FQ

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12N 15/85, 15/09 — A1

(11) International Publication Number: WO 92/20808

(43) International Publication Date: 26 November 1992 (26.11.92)

(21) International Application Number: PCT/US92/04054

(22) International Filing Date: 13 May 1992 (13.05.92)

(30) Priority data:
700,501 — 15 May 1991 (15.05.91) — US

(71) Applicant: CELL GENESYS, INC. [US/US]; 344 Lakeside Drive, Foster City, CA 94404 (US).

(72) Inventor: SMITHIES, Oliver ; 318 Umstead, Chapel Hill, NC 27514 (US).

(74) Agents: ROWLAND, Bertram, I. et al.; Flehr, Hohbach, Test, Albritton & Herbert, 4 Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, KR, LU (European patent), MC (European patent), NL (European patent), NO, SE (European patent).

Published
With international search report.
With amended claims.

(54) Title: GENOMIC MODIFICATIONS WITH HOMOLOGOUS DNA TARGETING

(57) Abstract

Methods are provided for modification of genomic target sites, where desired changes which can be small or subtle may be introduced into the target site to provide for modification of target genes or regulatory sequences. It is found that one may retain a marker without interference with the functioning of a target gene or select for excision of exogenous DNA to leave a single copy of the target gene with the modification.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/04054

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X – Y	Nature, V l. 336, issued 24 November 1988, Mansour et al., "Disruption f the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes", pages 348-352. See entire article.	12,13 ——— 5,9,10,14, 16, 17
Y	Proceedings of the National Academy of Sciences USA, Vol. 78, No. 10, issued October 1981, Orr-Weaver et al., "Yeast transformation: A model system for the study of recombination", pages 6354- 6358. See entire article.	2,3
Y	Nucleic Acids Research, Vol. 16 No. 18, issued 1988, Kim et al., "Recombinant fragment assay for gene targetting based on the polymerase chain reaction", pages 8887-8903. See entire article.	3
Y	Proceedings of the National Academy of Sciences USA, Vol. 76, No. 10, issued October 1979, Scherer et al., "Replacement of chromosome ssegments with altered DNA sequences constructed in vitro", pages 4951-4955. See entire article.	5,9,10,11
Y	Science, Vol. 242, issued 18 November 1988, Papayannopoulou et al., "Activation of developmentally mutated human globin genes by cell fusion", pages 1056-1058. See entire article.	11

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IE	Ireland	RO	Romania
CA	Canada	IT	Italy	RU	Russian Federation
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland			SN	Senegal
CI	Côte d'Ivoire	KR	Republic of Korea	SU	Soviet Union
CM	Cameroon	LI	Liechtenstein	TD	Chad
CS	Czechoslovakia	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

GENOMIC MODIFICATIONS WITH HOMOLOGOUS DNA TARGETING

ACKNOWLEDGEMENTS

This invention was supported in part by grants from NIH. The U.S. Government may have rights in this invention.

INTRODUCTION

Technical Field

The field of this invention is genomic modification using homologous DNA for targeting.

Background

There are a significant number of opportunities for introducing genetic modifications *in vivo* for purposes of correcting genetic defects and treating genetic disorders. Among genetic disorders involving hematopoietic cells are such defects as sickle cell anemia, β-thalassemia, various hemoglobinopathies, and disorders of erythrocyte metabolism including hereditary spherocytosis, pyruvate kinase deficiency, G6PD deficiency, etc. Among genetic disorders involving circulating plasma proteins and enzymes are inherited disorders of the complement system such as hereditary angioneurotic edema, agammaglobulinemia syndromes, α_1-antitrypsin deficiency and disorders of hemostasis, such as hemophilia A, hemophilia B and Von Willebrand's disease. Among gen tic disorders involving connective tissue, bone and muscle are the muscular dystrophies, the mucopolysaccharidosis syndromes, amyloidosis and various disorders of calcium and phosphate

metabolisms including hypophosphatasia, rickets and pseudo-hypoparathyroidism. Among the genetic disorders of metabolism are inherited disorders of amino acid metabolism, such as phenylketonuria, homocystinuria, albinism and tyrosinosis; inherited disorders of carbohydrate metabolism such as the glycogen storage diseases, diabetic syndromes and galactosemia; inherited disorders of lipid metabolism such as the hyperlipoproteinemias, the hyperlipidemias, the lipoprotein deficiency syndromes, the gangliosidoses, including Tay-Sach's disease, the lipidoses including Fabry's disease, Gaucher's disease, Refsum's disease, and Neimann-Pick disease; inherited disorders of steroid metabolism, such as adrenal hyperplasia; inherited disorders of purine and pyrimidine metabolism such as gout, Lesch-Nyhan Syndrome, and xanthinuria; and other metabolic disorders such as Wilson's disease, the porphyria syndromes, and hemochromatosis. Among the genetic disorders involving membrane transport of substances in the kidney, lung and other organs are the malabsorption syndromes, cystinuria, the renal tubular acidoses, cystinosis, Fanconi's syndrome and cystic fibrosis. Among disorders involving a genetic predisposition based on MHC antigen haplotype which may be potentially addressable by gene therapy are multiple sclerosis, ankylosing spondylitis, juvenile diabetes, rheumatoid arthritis and other autoimmune disorders.

Besides gene therapy, there may be an interest in modifying various domestic animals for a variety of purposes to improve their capabilities for use in supplying food, e.g., milk, butterfat, leaner meat, etc., to provide for animals which may be used for scientific investigations, e.g. Class I MHC deficient mice, and the like, and to produce proteins on a large scale, e.g., albumin and the like.

For these different interests, there will be different target cells for targeting for gene therapy or gene modification. Thus, one may be interested in modifying embryonal stem cells, somatic cells, hematopoietic stem cells, other stem cells, cells of connective tissue origin,

including myoblasts, osteoblasts, or chondroblasts; hepatic cells, endothelial cells, neural cells, epitithelial cells, and cells of endocrine origin, including islet cells, or the like. The techniques and methodology used for modifying the genotype of the target cells require that the modification provide the desired function. While one may rely upon random integration and selection of the clones of the integrants, selecting for function may not be sufficient. Random integration may result in a variety of situations, where the integrated DNA may be subject to regulation, depending upon the site at which it is integrated, where the functioning gene may become inactivated upon differentiation or proliferation of the cells, where the site of integration may result in a change in the functioning of one or more indigenous genes, or where the site of integration may lead to neoplasia.

In addition, it is not presently understood what the effect of having DNA, particularly foreign DNA, more particularly prokaryotic DNA, will have on the functioning of the gene being introduced, particularly, where one is interested in homologous recombination to achieve a modification of an indigenous gene. It is important to understand what the effect may be on the functioning capability of the DNA which is introduced as well as the target gene which is to be modified.

Relevant Literature

Valancius and Smithies, Mol. Cell. Biol. (1991) 11:1402-1408 describe an in-out targeting procedure for making genomic modifications in mouse embryonic stem cells. Hasty et al., Nature (1991) 350:243-246 describe the introduction of a mutation into the HOX-2.6 locus in embryonic stem cells by in-out targeting. Rothstein, Targeting, Disruption, Replacement and Allele Rescue; Integrative DNA Transformation in Yeast, Methods in Enzymology (1991); 281-301 describes in-out targeting in yeast. Smithies et al. Nature (1985) 317:230-234 describe targeting the human β-globin locus in a mouse

erythroleukemia hybrid cell line containing a single human chromosome eleven. Reports indicating that deletions or rearrangements involving the ends of targeting constructs and surrounding sequences sometimes accompany a gene targeting event may be found in articles by Doetschman, Maeda and Smithies, Proc. Natl. Acad. Sci. USA (1988) 85:8583-8587; Jason et al., Genes Dev. (1990) 4:157-166 and McMahon and Bradley, Cell (1990) 62:1073-1085. Also reported have been occasional secondary integrations accompanying an homologous recombination event, involving either the targeting constructed cells or cotransfected selectable DNA fragments. See Jason et al. and McMahon and Bradley, supra. See also Shulman et al. Mol. Cell. Biol. (1990) 10:4466-4472 and Stanton et al., ibid (1990) 10:6755-6758.

SUMMARY OF THE DISCLOSURE

Methods are provided for performing homologous recombination employing one or more selectable markers and a homologous region for introducing a modification in an indigenous chromosomal gene in a mammalian host cell. Two methods are employed for diminishing interference of the marker with the functioning of the target locus: (1) use of an Ω- (replacement) targeting vector, which allows retention of the selectable marker(s) in such a manner that the marker does not interfere with the desired function or expression of the indigenous gene; and (2) use of an O- (insertional) targeting vector, which allows for excision of the selectable marker(s). The resulting cells have the modified target locus without the indigenous sequence at the target site.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Methods are provided for targeted modification of indigenous genes using vectors comprising extended homology with the target gene, but differing at at least one site in conjunction with a marker gene. A linear DNA construct is employed with two regions of homology to the target locus.

Frequently, these regions will be proximal to the ends of the linear DNA molecule. The linear DNA molecule is transformed into the target cell by any convenient means and selection for integrants is performed using the selection provided by the marker gene. The selected clones are then further screened for target genes having the desired modification. If desired, one may further screen for loss of the selection gene and other foreign DNA with retention of the target gene with the introduced modification.

In targeting indigenous genes for modification, frequently subtle modifications, it is often important to ensure that the target locus is not modified in a way which interferes with the functioning of the target locus. When modifying indigenous DNA, one normally requires a selectable marker, which allows for selection of cells into which a construct has become integrated. The marker is normally integrated with the regions of homology and will be in close proximity to the region of homology. Therefore, methods must be devised which substantially ensure the functioning of the target locus.

Two different strategies are employed. In the first strategy an Ω-vector is employed, where hybridization results in a loop or D-loop of the hybrid. The resulting integrant retains the marker and provides a functioning target locus. The selectable marker is situated in such a manner in which at does not significantly interfere with the function or expression of the target locus. For example, the marker may be located 5' of the known local transcription regulatory sequences, within an intron or 3' of the coding region of the gene, etc. Usually the marker will be within 10 kbp, more usually within 5 kbp, of the gene.

In the second method, in-out targeting, an O-vector is employed where insertion results in two copies of the target locus homologous sequence: the indigenous sequence and the modified sequence. In a s cond step, an xcision step, the indigenous sequence and marker(s) are

excised leaving only the modified sequence. Thus, th end points are different for the integration and the excision.

The subject method can be used in a variety of ways for treating a variety of genetic diseases, mapping chromosomes, identifying loci, and the like. Of particular interest is the modification of dysfunctional genes, where the dysfunctional gene may be substituted with a functional gene. Thus, gene therapy may be carried out on a variety of types of cells, resulting in functional or modified genes from dysfunctional or undesired allelic genes. In addition, one may wish to modify a phenotype by modifying the capability of the functional gene, enhancing or diminishing the level of expression, changing the spectrum of activity of a pleiotropic gene, changing a particular allele, as in the case of major histocompatibility antigens, T-cell receptor variable regions, and the like.

The DNA employed for targeting will have a region of homology with the target locus differing from the locus by a modification, which may be a substitution, deletion, insertion, or combinations thereof. Also included will be a marker gene which allows for selection. One or more unique primer sites may be present for subsequent PCR analysis. The homologous DNA will usually be not more than about 100 kbp, usually not more than about 20 kbp and usually more than about 0.5 kbp.

The target cells may be any of a variety of vertebrate cells, particularly animals cells, more particularly mammalian cells, which may include any of the cells previously described. The constructs will comprise a region of homology associated with the target gene, where the region of homology may be noncoding, coding, or combinations thereof. A noncoding region may comprise the 5' non-coding region, introns, and in some instances the 3' noncoding region. The homologous region will normally encompass the modification, which may be a single site or a polynucleotide, usually not greater than about ten percent of the homologous r gion, e.g. 500 bp, usually not more than about five percent of the homologous r gion, e.g. 250 bp.

The modification will normally be bordered by a total of at 1 ast about 50 bp of homology, usually 100 bp and less than about 100 kbp usually less than about 50 kbp.

The restriction site which provides for the site of linearity of the DNA for the O-type vector employed for integration will be desirably between the site(s) of modification and the shorter stretch of homology. Desirably, one may have a short gap in the homology at the termini. The missing sequence which is filled in during the targeting can provide a primer site, so that targeted integration may be readily detected.

Depending on the nature of the vector, the organization of the functional sequences in the vector will vary. With the Ω-vector, the areas of homology will include regions flanking the marker(s) and the modified homologous region, where the flanking regions may or may not be immediately adjacent in the target locus. A double cross-over event is targeted resulting in replacement of the chromosomal region lying between the flanking homologous sequences. With the O-vector, the regions of homology to the chromosome are usually adjacent in the chromosomal target locus. The homologous region includes the modified region. Cross-over events resulting in the integration of the vector are selected, since the termini of the linear construct, when joined, define a sequence of substantially continuous homology. This results in the formation of two target loci, indigenous and modified, with the markers between the two target loci. Upon excision, the indigenous locus and markers will be lost and the desired modification will be retained, provided that the excisional cross-over occurs on the other side of the modification. The Ω-vector has an internal loop while the O-vector has an external loop of non-homologous sequence. Thus, in the Ω-vector the termini are distant at the target locus, while in the O-vector the termini are proximal at the target locus.

Various markers may be employed for selection. These markers include the HPRT minigene (Reid et al. (1990) Proc. Natl. Acad. Sci. USA 87:4299-4303, the neo gene for

resistance to G418, the HSV thymidine kinase (tk) gene for sensitivity to gancyclovir, the hygromycin resistance gene, etc. As indicated above for the O-vector, by linearizing within the region of homology, the marker gene(s), with accompanying foreign DNA (by foreign is intended foreign to the target host) will be situated between the duplicated genes, where one of the genes will have the introduced modification, while the other will be the indigenous sequence.

When carrying out in-out targeting, one may take advantage of using a marker that can be employed for both positive selection and negative selection of thee out step, e.g., hprt. Alternatively, one may use separate markers f r positive selection eg., neo, hygromycin resistance, etc., and for negative selection, eg., HSV-tk gene, cytosine deaminase, etc. The positive selection marker allows one to choose integrants lacking antibiotic resistance. Upon excision, the negative selection markers allows one to select against cells which retain the negative selection markers.

When carrying out targeting with an Ω-vector, one may employ a negative selection marker situated outside of the flanking homologous regions to enrich for double cross-over events.

Other aspects of the construct may include sequences which allow for specific primer regions for polymerase chain reaction (PCR) identification of homologous recombinants, one or more restriction sites, which allow for identification by gel electrophoresis, removal of a restriction site at the target locus, or other modification which allows for identification of target cells which have undergone the desired modification. In addition, the changes outside of the coding region should allow for retention of the transcriptional regulation region, unless some change in the transcriptional regulation region is desired. Therefore, the gene for s lection, restriction sites, primer sites, etc. will desirably be 5' or 3' of the coding region or within introns. The integrated DNA

sequence will usually be at least about 0.5 kbp, more usually at least about 1 kbp and usually less than about 100 kbp, more usually less than about 50 kbp.

Various techniques may be employed for introducing the linear DNA into the target cell. Techniques include electroporation, calcium precipitated DNA, fusion, transfection, and the like. The particular manner by which the DNA is introduced is not critical to this invention, although electroporation is preferred.

Once the target cells have been transformed, the cells may then be selected by means of the marker gene. Thus, the cells may be plated in a selective medium or grown in selective culture and clones identified for further investigation. Thus, where excision of the marker gene is not required, the clones may be analyzed using PCR, employing primers which will provide for different sized fragments, depending upon whether homologous recombination has occurred and whether the modified gene or the wild type gene is retained or other event has occurred to modify the target gene. In this way, target cells which have undergone the desired modification may be identified. Alternatively, one may look to the expression product by using antibodies specific for the modified gene expression product. Thus, one may perform any one of numerous immunoassays for identification of the expression product. Where the gene expresses a surface membrane protein, one may use monoclonal antibodies in conjunction with FACS for identification of cells expressing the desired product. It is found that the presence of the marker gene at the target locus does not significantly interfere with the expression of a target gene, allowing for substantially normal expression of the target gene in a host cell.

In some instances, one may wish to have the marker gene(s) removed along with the exogenous DNA. This may naturally occur as the result of various excision mechanisms, such as intrachromosomal recombination or homologous excision via unequal sister chromatidic exchange. In either event, a single copy of the gene will be obtained,

which will be either the indigenous gene or the modified gene.

As already indicated, the subject methodology may be used for gene therapy and mammalian fine-structure genetic analysis. Genes which may be targeted for gene therapy include β-globin, enzymes of erythrocyte metabolism, the complement system, coagulation factors, dystrophin, enzymes of carbohydrate, lipid, amino acid, steroid and purine and pyrimidine metabolism, transport proteins, e.g., cystic fibrosis transmembrane regulator, and the like.

The following examples are offered by way of illustration and not by way of limitation.

EXPERIMENTAL

Example 1. Correction of a Human β^s-Globin Gene By Use of a Replacement Vector.

Targeting Construct Construction and Preparation. The targeting construct β4.7NEO is a 4.7 kb *BamHI/Xba* I fragment that includes the β^A globin gene and surrounding sequences. It also contains a 20 bp oligomer inserted at the *Sph* I site 614 bp upstream of the start of the normal β globin transcript, and a 1.2 kb Xho I/Sal I fragment from pMC1neo Poly A (Thomas and Capecchi (1987) Cell 51:503-512) (the neomycin-resistance gene in this particular version of pMC1 neo Poly A, from Stratagene, contains a point mutation that reduces its ability to confer resistance to G418) inserted into a Bgl II site in the oligomer by blunt ending both oligomer and the insert DNA with Klenow polymerase.

For use in electroporations, the targeting sequences were excised from the vector plasmid Bluescript+ by a Sma I/Xba I double digest. This leaves one base of nonhomology at the 5' end of the targeting construct. The DNA was precipitated with EtOH after digestion and resuspended in phosphate buffered saline.

Cell Lines and Tissue Culture Conditions. The cell line BMS is a hybrid murine-human cell line derived from the fusion

of murine erythroleukemia cells (MEL 179, APRT-, available from Dr. A. Deisseroth) and human EBV-transformed lymphoblasts derived from an individual heterozygous for HPFH-2 and for the β^s globin gene. Fusion betwe n the two cell lines was carried out in the presence of PEG (polyethylene glycol) 15 (50% PEG 15 in 75 mM Hepes), and hybrid selection was achieved in AA media (50 μM Adenine, 40 μM Alanosine) in which only hybrid cells are expected to survive. After 2-3 weeks of selection, these hybrids appeared as clonal outgrowths and were tested for the presence of human chromosome 11 using a monoclonal antibody against a chromosome 11-encoded antigen (Papayannopolou et al., (1986) Cell 46:469-476). Hybrids were maintained in non-selective media and were occasionally enriched for the presence of human chromosome 11 by immunoadherence ("panning") to the monoclonal antibody. Cell line BSM carries only the copy of human chromosome 11 with the β^s allele, as judged by gene-specific polymerase chain reaction (PCR) amplification, and by determining the pattern of human globin expression after induction using previously reported methods (Papayannopolou et al., (1988) Science 242:1056-58)

Cell line PC4 was constructed as a positive control for PCR amplification with primers 1 and 2 (SEQ ID NOS: 1 and 2, respectively). The cell line contains a portion of the sequences in the targeting construct β4.7NEO (from the 3' end of oligomer used as a primer binding site for primer 2 (SEQ ID NO:2) through the 5' BamHI site), plus further upstream sequences from the globin region. PC4 therefore contains the binding sites for primers 1 and 2 (SEQ ID NOS: 1 and 2, respectively) and probes A and B, but lacks the neomycin gene and the β globin gene and therefore lacks binding sites for primers 3, 4, and 6 (SEQ ID NOS: 3, 4 and 6, respectively) and probe C. Clone PC4 was obtained by co-electroporating the construct into BSM cells with pMC1neo Poly A, followed by selection for G418 resistance and screening for PCR amplification of with primers 1 and 2 (SEQ ID NOS: 1 and 2, respectively).

All cells were grown at 37°C and 5% CO_2 in RPMI-1640 (Gibco) with 13% heat-inactivated fetal calf s rum, supplemented with 2 mM L-Glutamine. G418 selection was carried out with 300 μg/ml G418 sulfate (Gibco).

Electroporation. Cells were fed with fresh medium the day prior to electroporation, and were harvested when at a density of approximately 10^6 cells/ml. For electroporation, they were resuspended at $2x10^7$ cells/ml in warm growth medium, and digested. pβ4.7NEO was added to a final concentration of 5nM. Electroporation was with 10^7 cells in a chamber of 5 mm length and 100 mm^2 cross sectioned as described (Boggs et al., (1986) Exp. Hematol. 14:988-994). The electric pulse, from a 400 μF capacitor charged to 400V (800V/cm), was for one second. After electroporation, the cells were diluted into warm growth medium, and 10 ml was immediately plated into each of sixteen 100 mm diameter dishes at either $5x10^5$ or $1.5x10^5$ cells/dish. Electroporated cells were also plated into microtiter plates at 100 μl/well, at the same concentration and diluted threefold and tenfold. The next day an equal volume of medium containing 600 μg/ml G418 was added to the cultures. The number of G418-resistant clones in the 100 mm dishes was estimated by the Poisson distribution from the frequency of microtiter wells having no G418-resistant cells.

Probes. Probe A is a 218 bp Sty I/BamHI fragment from the genome just 5' of the targeting construct sequences. Probe B is a 627 bp Hpa I fragment from the 5' region of the targeting construct. Probe C is a 920 bp BamHI/EcoRI fragment covering the human β globin IVS2 region.

An RNA probe assaying human globin gene transcription was transcribed in vitro from a genomic 0.77 kb *EcoRI/Pst* I fragment cloned in the antisense orientation into a T7 vector; it contains the 3' end of the human β globin gene and has 212 bases of homology to the transcript. Th mouse probe was transcribed from a 0.65 kb *BamHI/HinfI* cDNA-derived fragment cloned into an Sp6 vector;

it has 298 bases of homology to the mouse β^{maj} gene transcript.

PCR. Polymerase chain reactions were as previously described (Kim and Smithies (1988) Nucleic Acids Res. 16:8887-8903) except that the $MgCl_2$ was 2mM. Screening of pools and individual clones was with 15 μl crude cell lysates, corresponding to approximately 10^4 cells, in a total reaction volume of 25 μl. Forty cycles were executed with denaturation for 1 minute at 90°C, and extension for 10 minutes at 60°C. Samples were analyzed by electrophoresis in 1.5% agarose gels followed by Southern blot analysis. Allele-specific PCR was for 30 cycles with 500ng of genomic DNA with primers and reaction conditions as described (Wu et al. (1989) Proc. Natl. Acad. Sci. USA 86:2757-2760). Samples were analyzed by electrophoresis in 6% acrylamide gels. Primer sequences are: primer 1 (SEQ ID NO:1) = 5'-CCCAGACACTCTTGCAGATT-3'; primer 2 (SEQ ID NO:2) = 5'-CAGATCTGGCTCGAGGCATG-3'; primer 3 (SEQ ID NO:3) = 5'-TGCGCTGACAGCCGGAACAC-3'; primer 4 (SEQ ID NO:4) = 5'-AATAGACCAATAGGCAGAG-3'; primer 5 (SEQ ID NO:5) = 5'-CACCTGACTCCTGT-3'; primer 6 (SEQ ID NO:6) = 5'-CACCTGACTCCTGA-3'. Note that primer 2 (SEQ ID NO:2) has 1 bp of nonhomology to the targeting construct at its 5' end.

Globin Analysis. The induction of globin synthesis and antibody labeling were as described (Papayannopolou et al., (1986) supra) with a β^s globin-specific monoclonal antibody (Papayannopolou et al., British J. of Hematology 35:25-31) and a more general human β globin-specific monoclonal antibody (Stammatayanopolous et al. (1983) Blood 61:530-539).

The slot blot hybridization assay used to assess relative levels of both murine and human globin cytoplasmic RNA in cells has been described (Constantoulakis et al. (1989) Blood 74:1963-1971).

Gen Targeting. Gene targeting was used to correct the β^s globin gene on a human chromosome 11 in mouse erythroleukemia hybrid cell lines BSM. The β^A globin replacement (or Ω) type targeting construct, β4.7NEO, has 4.7 kb of sequences homologous to the human β^A globin region; it also contains a unique oligomer for use as a PCR primer binding site (primer 2) (SEQ ID NO:2), and a neomycin-resistance gene, both placed 5' of any known local β globin transcriptional regulation sites. The neomycin gene is a 1.2 kb fragment derived from pMC1neo PolyA and is driven by the herpes simplex thymidine kinase (tk) promoter plus the duplicated mutant polyoma virus enhancer originally designed for use in mouse embryonic stem cells (Thomas and Capecchi (1987), supra).

The targeting construct was introduced into BSM cells in a series of eight electroporations. After electroporation, the cells were diluted, and a total of $4.1x10^7$ were immediately plated into tissue culture dishes. The following day they were placed under G418 selection which yielded 126 pools of between 10 and 1000 G418-resistant clones per dish (average about 200), each clone having incorporated the targeting construct somewhere in the genome.

Detection of Targeted Clones. Upon completion of selection, and growth to adequate density (10 to 20 days), a small portion of each pool was removed and tested by a PCR assay (Kim and Smithies (1988) Nucleic Acids Res. 16:8887-8903) for the presence of a targeted clone. Two primers were used: primer 1 (SE ID NO:1) is specific to the target locus, since it is from approximately 400 bp upstream of the 5' end of the targeting fragment; primer 2 (SEQ ID NO:2) is specific to the incoming DNA, since it is from the synthetic oligomer sequences present only in the targeting DNA. Specific PCR amplification to give a 1.2 kb diagnostic band hybridizing to probe A can only occur if a targeting event juxtaposes these two primers. The PCR results led to the identification of a pool containing a targeted clone.

Although the PCR generated two hybridizing bands, rather than the expected single band, the pattern is still indicative of a targeted recombinant since the positive control cell line PC4 yields the same two bands, while the parental BSM cells yield neither. In all, a total of three PCR positive pools were identified among the 126 pools examined. This corresponds to a targeting frequency of one targeted clone in about 9,700 G418-resistant clones.

The PCR analysis included positive control lysates made from mixtures of PC4 cells with untreated BSM cells at ratios of 1 to 10 and 1 to 100. PC4 is a pseudo-recombinant cell line which contains integrated copies of foreign DNA having primers 1 and 2 (SEQ ID NOS: 1 and 2, respectively) already juxtaposed. This positive control proved essential for working out appropriate PCR conditions, but introduced the risk that false positives might arise from contamination of nontargeted cells with the diagnostic 1.2 kb fragment from amplified controls. The possibility that the positives were artifactual contaminants of this type was excluded by carrying out a second set of PCR amplifications using a neomycin gene-derived primer (primer 3) (SEQ ID NO:3) in place of primer 2 (SEQ ID NO:2). All three pools that the test PCR assay had indicated contained a targeted clon yielded the expected 1.6 kb band that hybridized to probe A during this second PCR reaction; as expected, the positive control PC4 cells did not.

Isolation of a Targeted Clone. Sib-selection was used to isolate a clone of targeted cells from one of the PCR-positive pools that contained about 200 independent G418-resistant clones. Cells from this pool were diluted into 96 smaller pools with approximately 10 cells in each. After expansion, these smaller pools were rescreened by PCR for the presence of the diagnostic 1.2 kb band. Three of the smaller pools gave a positive PCR result, and the signal level corresponded to that expected for mixtures having one targeted cell for every t n nontargetted.

One of the enriched pools was then diluted into microtiter dishes so that only 10 to 20% of the wells received a cell. After growth to adequate density, samples were removed from 66 wells containing growing cells from PCR amplification. Three of the 66 clone wells showed the 1.2 kb diagnostic band at the same level as undiluted PC4 control cells, and were therefore presumed to contain clones of targeted cells.

Southern Blot Analysis. One isolate of the presumed targeted clone was expanded, and used to prepare genomic DNA for Southern blot analysis. When the 5' probe B was used, a 12.8 kb band generated by *Pvu* II digestion of genomic DNA from the starting BSM cells is replaced by a predicted 6.9 kb in the targeted cells; likewise a parental 5.5 *Eco*RI fragment is replaced by a predicted 3.7 kb band. Similarly, with the 3' probe C, a parental 5.0 kb *Bgl* II band becomes, as predicted, 4.0 kb after targeting, and, again as predicted, the parental 12.8 kb *Pvu* II band becomes the predicted 7.1 kb band.

Southern blot analysis was also used to investigate the fidelity of the targeting event. A *Bam*HI site that begins 1 bp from the 5' end of the targeting construct is still intact after targeting, as shown by the presence of a predicted 2 kb fragment that hybridizes to probe B following *Bam*HI digestion of DNA from the targeted clone; the parental BSM DNA gives a 1.9 kb band. Likewise, the 3' *Xba* I site at the 3' end of the targeting construct was shown to be intact by the presence of a 4.5 kb band that hybridized to probe C following an *Xba* I/*Pvu* II double digest of the targeted clone, compared to the 10 kb band from parental BSM DNA. These hybridization patterns establish that no end deletions have occurred and also re-confirm the targeting event itself.

The Southern blots, and comparable blots using neomycin- and vector-specific probes from the targeting plasmid, show that the genome of the targeted clone contains

only a single copy of the targeting construct integrated at the desired location.

Confirmation of Correction to the β^A allele. Two independent methods were used to demonstrate that the β^S allele had been changed to β^A by the targeting. First, allele-specific PCR was performed on various cells using primer 4 (SEQ ID NO:4) together with either the β^S-specific primer 5 (SEQ ID NO:5), or the β^A-specific primer 6 (SEQ ID NO:6). The parental (β^S-containing) BSM cells, as expected, yielded no amplified band when the β^A-specific primer 6 (SEQ ID NO:6) was used, but a band was obtained with the β^S-specific primer 5 (SEQ ID NO:5). A pool of G418-resistant, but nontargetted cells yielded an amplified band when either primer set was used, although a more intense band was observed with the β^A-specific primer, than with the β^S-specific primer. This result is expected because the human chromosome 11 is not present in all of the hybrid cells, whereas a randomly integrated targeting construct must be present in all cells (in order to obtain G418 resistance), and often occurs in multiple copies. The targeted clone, again as expected, amplifies a band when the β^A-specific primer 6 (SEQ ID NO:6) is used, but not with the β^S-specific primer 5 (SEQ ID NO:5). These observations establish that the targeted clone now contains a β^A globin gene, but no longer carries the β^S gene.

Secondly, allele-specific antibodies were used to investigate the globin polypeptides synthesized after the parental BSM cells and the targeted clone have been induced with HMBA (hexamethylenebisacetamide). An antibody specific for human β globin, but unable to discriminate between β^S and β^A globin, shows the presence of some type of human β globin in both the induced BSM cells and the induced targeted clone. In contrast, a human β^S-specific antibody binds to the globins produced by the induced parental BSM cells, but not by the targeted clone. These observations establish that although the targeted clone is able to

synthesize human β globin, it can no longer synthesize β^s globin.

Regulation of Globin Expression. A slot blot hybridization assay with probes specific for murine β^{maj} globin and human β globin was used to ask whether or not the targeted clone had its ability to undergo human β^A globin induction altered. Specifically the ratios of induced to uninduced cytoplasmic RNA for both murine β^{maj} and human β globin transcripts within cell line BSM was determined and found to be essentially identical (17:1 and 15:1 respectively). Similarly, the ratios of induced to uninduced levels of murine β^{maj} compared to human β globin within the targeted clone were also essentially identical (7:1 and 5:1 respectively). Since the inducibility of the murine and human β globin genes are essentially the same within the polyclonal parental BSM cells, and within the targeted clone, we conclude that the targeting event has not significantly altered the ability of the (now corrected) human gene to respond to induction.

The above results demonstrate a frequency of targeting of at least one targeted clone for 9,700 G418-resistant clones. The Southern Blot analysis rigorously established that the isolated clone had been targeted as planned and verified the fidelity of the targeting event in the isolated clone. No detectable secondary events occurred along with the gene targeting event. Human globin genes introduced on their native human chromosomes introduced into the BSM cells by somatic cell fusion are regulated and expressed after induction in a manner comparable to that shown by the endogenous mouse globin genes (Marks and Rifkind (1978) in Ann. Rev. Biochem. 47:419-448; Willing et al. (1979) Nature 277:534-538; Deisseroth and Hendrick (1979) Proc. Natl. Acad. Sci. USA 76:2185-2189. The results demonstrate that gene targeting can correct a human β^s globin gene to β^A, an essential requirement before gene targeting can be considered for human gene therapy. Furthermore, the induction ratio of the corrected gene was

not significantly altered by the introduction of a neomycin-resistance helper gene into the target locus to facilitate identification and isolation of the targeted clone.

Example 2. The Use of In-Out for Making Subtle Genomic Modifications in Mouse Embryonic Stem Cells.

Cell culture. The mouse ES cell line E-14TG2a was isolated as described previously (Hooper et al. (1987) Nature 326:292-295; Thomas and Capecchi (1987), supra). Cells were grown in Dulbecco's modified Eagle's medium (GIBCO) supplemented with 15% heat-inactivated fetal calf serum (Flow) and 10 μM 2-mercaptoethanol (Sigma). The pluripotential nature of the ES cells was retained by supplementing each liter of growth medium with 10^6 U of recombinant human leukemia inhibitory factor (available from N. Gough, Walter and Eliza Hall Institute, Melbourne, Victoria, Australia). Because feeder layers were not used, all culture dishes were coated with 0.1% sterile gelatin to ensure cell adhesion. HAT medium was standard culture medium supplemented with 120 μM hypoxanthine, 0.4 μM aminopterin, and 20 μM thymidine. 6-TG (thioguanosine) selection was carried out in standard medium containing 10 μM 6-TG. Cultures were incubated at 37°C in an atmosphere of 5% CO_2. They were checked periodically for mycoplasma contamination.

Vectors. Plasmid pNMR133 has already been described (Doetschman et al., (1987) J. Embryol. Exp. Morphol. 330:576-578). It contains 5 kb of DNA identical to the exon 3 target region of the mouse HPRT gene, except for a 4-bp insertion that destroys a unique HindIII site and consequently generates a new NheI site in intron 2. It also carries the human HPRT promoter and exon 1 sequences (which have been shown to function in mouse cells) and the mouse exon 2 region.

Plasmid pNMR133D200 was derived from pNMR133 by removing a 200-bp BglII fragment from intron 2.

DNA preparation. Targeting vector DNAs were prepared by standard methods, omitting the CsCl purification, which was found unnecessary. All targeting DNAs were linearized by restriction enzyme digestion, using the manufacturers' recommended conditions, prior to electroporations. Digested DNAs were ethanol precipitated and resuspended in sterile TE buffer (0.05 M Tris, 0.001 M EDTA).

DNA transfers and selections. The vectors were introduced into the ES cells by electroporation (Boggs et al. (1986) Exp. Hematol. 14:988-994). The cells were grown in 100-mm culture dishes (as described above) to a density of 1 x 10^7 to 2 x 10^7 cells per dish in nonselective medium. Cultures were trypsinized, centrifuged, and then resuspended in nonselective medium to a density of 4 x 10^7 to 10 x 10^7 cells per ml. A 0.5-ml sample of the cell suspension was added to each microfuge tube, and prepared DNA was then added to a final concentration of 5 nM. The cell-DNA mixtures were incubated on ice for 20 min, loaded into an electroporation chamber precooled on ice (length, 5 mm; cross section, 100 mm^2), and exposed to a 1-s electrical pulse from a 250-μF capacitor charged to 300 V. Cells were immediately removed from the chamber and plated into five 100-mm cultures dishes. The plates had been prepared by gelatinization and contained 7 ml of nonselective medium. The cells were allowed to recover overnight. The next day, the number of colonies in each dish was determined by counting, and HAT selection was then applied.

Cultures to be assayed for the loss of HPRT function by selection in 6-TG were maintained under HAT selection for at least 1 month prior to the start of the assay in order to kill any accumulated hprt$^-$ cells. These cultures were trypsinized, counted, and then replated at a density of 0.5 x 10^7 to 1 x 10^7 cells per plate in nonselective medium. They were without selection for 3 or 4 days to allow spontaneous revertants time to purge residual HPRT transcripts or protein. Selection was then started by applying 6-TG medium.

All selections were maintained for 16 days, with feeding as necessary. Targeting and reversion frequencies were determined by counting the number of resistant colonies obtained for each experiment. Individual colonies were picked by using cloning rings into 24-well (1 ml per well) dishes and maintained under selection. Cultures were transferred to 60-mm culture dishes and then either harvested for genomic DNA preparation or transferred to 100-mm dishes for further expansion.

Genomic DNA preparation and characterization. DNA was prepared from expanded clones by using conventional procedures. Restriction enzyme digestions were done according to manufacturers' specifications, incubating overnight. After electrophoresis on 0.8% agarose gels, Southern blotting was done by standard techniques.

Probes. Two probes were used, a 250-bp RsaI fragment from intron 3 and a 300-bp HindIII-XhoI fragment from the human cDNA which includes exons 3 to 6 but is specific for the mouse exon 3 element (Doetschman et al., (1987) Nature 330:576-578). Both probes hybridize to sequences present in the endogenous locus as well as on the targeting vectors. For each blot, 25 to 50 ng of purified fragment was radiolabeled with ^{32}P-dCTP by the random-primed oligonucleotide method, using a Boehringer Mannheim kit. Four-hour prehybridizations and overnight hybridizations were done in 50% formamide solutions at 42°C. Blots were washed to a stringency of 1 x SSC (0.15M NaCl plus 0.015 M sodium citrate) at 68°C. Washed blots were exposed to preflashed XAR-5 film at -70°C.

In step: the integration event. The first step in the two-step targeting procedure is a homologous integration event that incorporates vector DNA carrying the desired modification into the genome. The method of (Doetschman et al. (1987) supra), was used to introduce into mouse ES cells an integrating targeting vector that carries a 4-bp

insertion in the second intron of the HPRT gene. Either plasmid pNMR133 or plasmid pNMR133D200 (which has a 200-bp gap in the region homologous to the target locus) was electroporated into the male mouse-derived ES cell line E-14TG2a. This cell line, isolated by Hooper et al. ((1987) supra), as a spontaneous mutation in culture, carries a nonreverting deletion of the promoter and first two exons of the nine-exon, 33-kbp, X-linked HPRT gene (Thompson et al., (1989) Cell 56:313-321) rendering it phenotypically hprt$^-$. Both targeting vectors contain approximately 5kb of DNA identical in sequence to the exon 3 target region of the hprt$^-$ gene except for the intended modification: a 4-bp insertion in intron 2 that destroys a unique HindIII site. They also carry the human HPRT promoter and exon 1 sequences and the mouse exon 2 region.

The homologous integration event generates a duplication of the 5-kb target region separated by the remainder of the vector sequences. The duplicated regions are identifiable with the exception of the 4-bp insertion, identified by a missing HindIII site, that is located on the downstream repeat. This event restores the promoter and first two exons deleted from the locus, generating HPRT+ targeted recombinants that can be directly selected with HAT-containing medium.

Three independent HPRT+ cell lines were isolated by selection in HAT medium at an average frequency of 2.8 x 10^{-6} per electroporated cell and it was then confirmed that these cell lines were targeted by genomic Southern blot hybridization. The blots were probed either with a 250-bp RsaI fragment from intron 3 or with a 300-bp HindIII-XhoI fragment from the human cDNA that specifically hybridizes to the mouse exon 3. Both probes hybridize to sequences found in the genome as well as on the targeting vectors. All of the cell lines examined contained the expected recombinant locus, indicating that a single copy of the targeting vectors had integrated into the E-14TG2a hprt$^-$ gene.

Cell lines A and C hybridized to the single 19kb HindIII fragment expected for a simple insertion of the

12-kb vector into the 7 kb endogenous fragment. Cell line D hybridized to two HindIII fragments, the endogenous 7-kb and the vector 12-kb fragments. This cell line, generated with plasmid pNMR133D200, has lost the 4-bp insertion as a consequence of the integration event, so that revertants obtained from this line could not be properly modified. However, it was used in the excision experiments (see below) since it could still generate useful information about the frequency and accuracy of the excision reaction. BamHI digestion of all recombinants revealed the expected 9.4-kb endogenous band and the 6.9-kb vector-derived band. No extraneous bands could be detected, confirming that all of the recombinants carried single-site, single-copy insertions of the vector DNAs. In addition to these three lines, one more cell line, B, generated previously in the laboratory (Doetschman et al., (1987) supra) was used in the excision studies (see below). This line carries the same recombinant locus found in lines A and C.

Out step: the excision event. The second in the two-step targeting procedure is a spontaneous event that excises from the genome the vector sequences that integrated in the first step. A homologous recombination event between the regions duplicated during the in reaction can occur by either intrachromatid recombination (Doetschman et al., (1987) supra) or unequal sister chromatid exchange. A crossover event in the 2-kb region comprising the 5'-terminal portion and the HindIII site will leave the 4-bp insertion in the genome; crossing over in the 3 kb region comprising the 3'-terminal portion will excise the 4-bp modification along with the vector sequences or move it to the (HPRT+) triplicated chromosome. Either way, the excision event removes the vector-derived promoter and first two exons, causing a reversion to the hprt⁻ phenotype. Such revertants can be selected with the nucleoside analog 6-TG.

The four HAT cell lines described above were used to study the excision (out) reaction. They all carry essentially the same HPRT locus: a duplication of 5 kb

separated by 7 kb of plasmid-derived unique sequence. ES cell line D-3 (Doetschman (1985) J. Embryol. Exp. Morphol. 87:27-45), which carries the wild-type HPRT locus, was used as a control in these experiments to determine the spontaneous rate of mutation from HPRT+ to hprt⁻ at the normal locus. The experiments were performed as described above. The day after replating, the number of colonies observable in each dish was determined by counting. Typically, ES cells form one colony for every 5 to 10 cells plated (Table 1).

TABLE 1
Frequency of the Out Reaction

Cell Line	Cells Plated (10^7)	Colonies 25 h Post (10^6)[a]	6-TG^r Colonies[b]	Reversion Frequency x 10^{-7}
A	2.90	3.2	23	7.9
B	2.2	3.2	13	5.9
C	4.2	6.4	14	3.3
D	3.8	4.0	56	14.7
Total	13.1		106	8.1
D-3 (control)	5.6	4.8	0	<.18

[a] Number of colonies counted the day after replating, reflecting a 10 to 20% plating efficiency.

[b] Number of colonies counted after 2 weeks of 6-TG selection.

[c] Number of 6-TG^r colonies obtained per plated cell.

This is due to their propensity to form aggregates, not to a high death rate. That is to say, each colony found the day after replating is composed of 5 to 10 individual cells. Although this aggregation may interfere with the 6-TG selections as a result of metabolic cross-feeding (Hooper et al., (1981) Int. Rev. Cyto. 69:45-104), it cannot be avoided.

The number of 6-TG^r colonies obtained for each line examined and calculated reversion frequencies are listed in Table 1. As shown, all four HAT lines initially generated by gene targeting reverted to the hprt⁻ phenotype at similar frequencies, averaging 8 x 10^{-7} 6-TG^r colonies

isolated for every HAT^r cell plated. Control cell line D-3, which carries the wild-type HPRT locus, failed to produce any 6-TG^r colonies from 5.6 x 10^7 cells plated. Thus, the spontaneous mutation frequency at the HPRT locus, for cells preselected with HAT, is less than 1.8 x 10^{-8}. This result is consistent with the rate of 1.5 x 10^{-8} per cell generation reported for the locus by Caskey and Kruh, (1979) Cell 16:1-19.

Several of the individual 6-TG^r colonies were analyzed further by genomic Southern blot hybridization. Using the HPRT-specific probes described above. The number of colonies examined from each line and a summary of the results obtained from the Southern blot hybridizations are presented in Table 2.

TABLE 2

Accuracy of the Out Reaction

Cell Line	Colonies examined	Accurate revertants[a]	Revertants with 4-bp insert[b]
A	3	2	2/2
B	9	7	6/7
C	11	11	11/11
D	3	3	NA
Total	26	23	19/20

[a] Number of colonies containing the expected hprt$^-$ locus.

[b] Number of hprt$^-$ colonies that retain the 4-bp insertion presented as a fraction of the number of accurate revertants obtained.

NA. Not applicable: this cell line does not carry the 4-bp insertion.

Of a total of 26 6-TG^r colonies examined, 23 (88%) had executed the out reaction and accurately excised the

integrated vector sequences from the genome, as determined by the genomic Southern blots. They all revealed a single 9.4-kb BamHI band upon hybridization, the size predicted for a simple homologous excision event. This is the same BamHI fragment found in the parental E-14TG2a hprt$^-$ locus. HindIII digestion of the revertant DNAs is expected to reveal one of two bands upon hybridization: either an 11-kb fragment, if the crossover occurs in the 5' region and the 4-bp insertion introduced by the in event is retained, or a 7-kb fragment, if the crossover occurs in the 3' region and the modification is removed from the hprt$^-$ genome.

Of the 23 out revertants examined, 20 were derived from HPRT+ cell lines that carried the 4-bp insertion initially introduced by the targeted integration event; 19 of these colonies contain the single 11-kb HindIII fragment, indicative of the accurate excision event which retains the 4-bp insertion. Thus, these 19 colonies have been correctly modified by the in-out targeting procedure. Only 1 of these 20 revertant colonies lost the 4-bp modification, as determined by the presence of a 7-kb HindIII fragment. Therefore, 95% of the accurate revertants which could have retained the 4-bp insertion did so. The three remaining colonies which were found to have been generated by the out reaction were derived from the HAT cell line D. As this cell line does not carry the 4-bp modification, the revertants revealed only the 7-kb HindIII band upon hybridization.

To confirm that the 11-kb HindIII band characterizing the accurately modified hprt$^-$ revertants retained the 4-bp insertion initially introduced by the targeting vector, two of the genomic DNAs were digested with NheI site. The 4-bp insertion introduced to destroy the HindIII site in the original targeting vector generates a unique NheI site. In the case of the revertants that have retained the 4-bp insertion (11-kb HindIII), NheI digestion will generate a 2.7-kb band which hybridizes to the probes. In the case of the revertants which have lost the insertion

(7-kb HindIII), a 4.9-kb band will result. The accurately modified revertant does contain a 2.7-kb NheI fragment which hybridizes to the probe, confirming the presence of the 4-bp insertion, and the revertant which has lost the 4-bp insertion reveals a 4.9-kb fragment upon hybridization.

The other three colonies examined were found to contain aberrant hprt$^-$ loci that did not arise by the predicted homologous excision reaction. They contained a single 14-kb HindIII fragment and a 16-kb BamHI fragment that hybridized to the probes. These fragments failed to hybridize to a plasmid-specific probe, indicating that the target vector sequences have been excised from the genome. Since the bands are not the expected sizes, these colonies were probably generated by an alternate excision reaction. Because they account for only 12% of the 6-TG^r colonies obtained, they were not examined further.

The above data demonstrate the successful in-out targeting in modifying the genome of a mouse ES cell line by introducing a 4-bp insertion, in a two step procedure, the second step being automatic. The average frequency of the in reaction was found to be 2.8×10^{-6}. The frequency of the out reaction, 8×10^{-7} per HAT^r cell plated is approximately 30% that of the in reaction. This frequency is 40-fold higher than the spontaneous mutation rate at the normal HPRT locus. Of the 6-TG^r colonies isolated, 88% had accurately excised the target vector sequences from the genome.

It has been shown that metabolic cross-feeding can interfere with 6-TG selections; therefore, the reversion frequency determined may be an underestimate. Because the ES cells always form aggregates upon plating, the possibility of such cross-feeding could not be eliminated. This suggests that when using the in-out targeting procedure in other cell lines, it may be useful to plate HAT^r cells at a lower density to minimize the potential of revertant loss due to such cross-feeding.

By not growing the ES cells on feeder layers, the pluripotential nature of the cells was retained. This was achieved by supplementing the growth medium with purified

human leukemia inhibitory factor. (Smith et.al., (1988) Nature 336:688-690; Williams et.al., (1988) Nature 336:684-687). This modification greatly simplifies the 2-step selection procedure.

The above results show that both the integration and excision events can occur accurately and with a frequency sufficient for use in a 2-step targeting technique. While the above procedure employs the directly selectable the HPRT locus, the same procedure could be adaptable to modify non-selectable loci in an hprt$^-$ cell line by using the HPRT minigene described by Reid et.al., (1990) Proc. Natl. Acad. Sci. USA 87:4299-4303. The minigene would be carried on an integrating targeting vector, thereby allowing selection to be used for both the integration and excision events. Homologous recombinants are likely to be found after the in step at a frequency of 1 in 1,000 HAT^r cells, this being the ratio of transformed to targeted cells reported previously. The targeted cell lines can then be identified by the polymerase chain reaction, for example, or other means, depending upon the nature of the targeting gene and the modification. Including a small gap in the region of homology on the insertional vector provides a convenient primer binding site, since all gaps are repaired during the homologous insertion event. Excision-derived hprt$^-$ revertants are likely to be found after the outstep at a frequency of nearly 1 in 10^6 per targeted cell line.

The above techniques allow for modification of genomes of viable cells, particularly embryonic cells, where a stable modification can be achieved, which can be inherited by progeny cells. In addition, the modifications can be subtle and functional target genes achieved even when a marker is allowed to remain in the genome. Thus, the subject invention demonstrates the feasibility of gene therapy with stem cells or other cells, which can be used for the treatment of a variety of genetic or other diseases.

All publications and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or

patent application was specifically and individually indicated to be incorporated by reference.

The invention now being fully described, it will be apparent to one of ordinary skill in the art that many changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Smithies, Oliver

(ii) TITLE OF INVENTION: Genomic modifications with homologous DNA targeting.

(iii) NUMBER OF SEQUENCES: 6

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Bertram I. Rowland
(B) STREET: 4 Embarcadero Center, Suite 3400
(C) CITY: San Francisco
(D) STATE: CA
(E) COUNTRY: USA
(F) ZIP: 94111

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US 07/700,501
(B) FILING DATE: 15-MAY-1991
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Rowland, Bertram I
(B) REGISTRATION NUMBER: 20,015
(C) REFERENCE/DOCKET NUMBER: A55082/BIR

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 781-1989
(B) TELEFAX: (415) 398-3249

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCAGACACT CTTGCAGATT 20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: singl
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAGATCTGGC TCGAGGCATG 20

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TGCGCTGACA GCCGGAACAC 20

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

AATAGACCAA TAGGCAGAG 19

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CACCTGACTC CTGT 14

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CACCTGACTC CTGA 14

WHAT IS CLAIMED IS:

1. A method for introducing changes at a target locus in a chromosome of a viable mammalian cell, said method comprising;

transforming said cell with a linear DNA construct comprising a sequence having at least 50 bp of homology with an indigenous region of said target locus, said homology comprising a sequence different from said target locus, and a marker gene allowing for selection of cells comprising said marker gene, wherein said DNA construct is an Ω-targeting vector or an O-targeting vector, wherein a non-homologous sequence forms an internal loop or an external loop, respectively;

growing said transformed cells in selective medium to provide marker gene containing cells; and

isolating cells comprising said change in said indigenous region by identifying the presence of said construct sequence at said locus, wherein when said construct is an Ω-vector said non-homologous region is at a site which does not substantially interfere with the functioning of said target locus.

2. A method according to Claim 1, wherein said linear DNA construct is an O-vector and is linearized in the homologous sequence and has a gap in homology with said indigenous region, when the ends of said linear sequence are joined.

3. A method according to Claim 2, wherein said gap defines a primer sequence, and the polymerase chain reaction is used to identify said subtle change by using a primer complementary to the homologous sequence at the gap.

4. A method according to Claim 1, wherein said transforming is by electroporation.

5. A method according to Claim 1, wherein said vector is an O-vector and including the additional step of growing said isolated cells for sufficient time for the indigenous region and markers of said target locus to be excised; and

identifying cells comprising the homologous sequence and lacking said indigenous sequence, but retaining said modification, by means of a marker allowing for negative selection.

6. A method according to Claim 1, wherein said mammalian cells are embryonic cells.

7. A method according to Claim 1, wherein at least one terminus of said linear DNA construct comprises at least about 50 bp of homology with said indigenous region.

8. A method according to Claim 1, wherein said target locus comprises a defective globin gene and said construct comprises a functional globin gene.

9. A method for introducing a change in a gene at a target locus in a chromosome of a viable mammalian cell, said method comprising;

transforming said cell with a linear DNA construct comprising an O-targeting vector having (1) a sequence of at least 50 bp of homology with an indigenous region of said gene and differing from said indigenous region and (2) one or a combination of marker genes allowing for positive and negative selection of cells comprising said marker gene(s);

growing said transformed cells in selective medium to positively select for marker gene containing cells in a first step and negatively select for marker gene containing cells in a second step; and

isolating cells comprising said subtle change in said gene.

10. A method according to Claim 9, wher in said construct comprises a hprt gene.

11. A method according to Claim 9, including identifying said modified cells, wherein said identifying is by means of monoclonal antibodies specific for the polypeptide encoded by said homologous sequence.

12. A method for introducing a change in a gene at a target locus in a chromosome of a viable mammalian cell, said method comprising;

transforming said cell with a linear DNA construct comprising an Ω-targeting vector having (1) a sequence of at least 50 bp of homology with an indigenous region of said gene and differing from said indigenous region and (2) at least one marker gene allowing for positive selection of cells comprising said marker gene;

growing said transformed cells in selective medium to positively select for marker gene containing cells; and

isolating cells comprising said change in said gene.

13. A method according to Claim 12, wherein said marker gene is antibiotic resistance.

14. A method according to Claim 12, wherein said marker gene is 5' of said gene and flanked 5' by at least 50 bp of homologous sequence.

15. A linear targeting vector construct comprising a wild-type structural gene sequence of a gene commonly associated with a genetic disease as a result of a difference in sequence, said wild-type structural gene being homologous at the chromosomal locus of said difference, at least one marker gene for positive selection and having flanking homologous sequences to said marker gene, wherein the homologous sequences proximal to the termini of said vector construct are either distal at said chromosomal locus to define an

Ω-targeting vector or proximal at said chromosomal locus to define an O-targeting vector.

16. A targeting vector according to Claim 15, wherein said at least one marker comprises a marker for positive selection and a marker for negative selection.

17. A targeting vector according to Claim 15, wherein said vector is an Ω-targeting vector and comprises a marker gene at one terminus.

18. A targeting vector according to Claim 15, wherein said gene is the β-globin gene.

AMENDED CLAIMS

[received by the International Bureau
on 18 September 1992 (18.09.92);
original claims 1,5,9-11 and 14 amended;
remaining claims unchanged (3 pages)]

5. A method according to Claim 1, wherein said vector is an O-vector and including the additional step of growing said isolated cells for sufficient time for the indigenous region and markers of said target locus to be excised; and

identifying cells comprising the homologous sequence and lacking said indigenous sequence, but retaining said subtle change, by means of a marker allowing for negative selection.

6. A method according to Claim 1, wherein said mammalian cells are embryonic cells.

7. A method according to Claim 1, wherein at least one terminus of said linear DNA construct comprises at least about 50 bp of homology with said indigenous region.

8. A method according to Claim 1, wherein said target locus comprises a defective globin gene and said construct comprises a functional globin gene.

9. A method for introducing a subtle change in a gene at a target locus in a chromosome of a viable mammalian cell, said method comprising:

transforming said cell with a linear DNA construct comprising an O-targeting vector having (1) a sequence of at least 50 bp of homology with an indigenous region of said gene and differing from said indigenous region and (2) one or a combination of marker genes allowing for positive and negative sel cti n of cells comprising said marker gene(s);

using said marker g ne for s lection by growing said transformed cells in sel ctive medium to positively s lect for marker gene c ntaining cells in a first step and negatively sel ct for marker gene containing c lls in a s cond step; and

isolating cells comprising said subtle change in said gene.

10. A method according to Claim 9, including identifying said cells comprising said subtly changed gene, wherein said identifying is by means of monoclonal antibodies specific for the polypeptide encoded by said homologous sequence.

11. A method for introducing a subtle change in a gene at a target locus in a chromosome of a viable mammalian cell, said method comprising:

transforming said cells with a linear DNA construct comprising an Ω-targeting vector having (1) a sequence of at least 50 bp of homology with an indigenous region of said gene and differing from said indigenous region and (2) at least one marker gene allowing for positive selection of cells comprising said marker gene;

growing said transformed cells in selective medium to positively select for marker gene containing cells; and

isolating cells comprising said subtle change in said gene.

12. A method according to Claim 11, wherein said marker gene is antibiotic resistance.

13. A method according to Claim 11, wherein said marker gene is 5' of said gene and flanked 5' by at least 50 bp of homologous sequence.

14. A linear targeting vector construct comprising a wild-type structural gene sequence of a gene commonly associated with a genetic disease as a result of a difference in sequence, said wild-type structural gene being homologous at the chromosomal locus of said difference, at least one marker gene for positive selection and having flanking homologous sequences to

said mark r gene, wherein the homologous sequences proximal to the termini of said vector construct are eith r distal at said chromosomal locus to define an Ω-targeting vector or proximal at said chromosomal locus to define an O-targeting vector.

15. A targeting vector according to Claim 14, wherein at least one marker comprises a marker for positive selection and a marker for negative selection.

16. A targeting vector according to Claim 14, wherein said vector in an Ω-targeting vector and comprises a marker gene at one terminus.

17. A targeting vector according to Claim 14, wherein said target gene is the β-globin gene.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US92/04054

A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) :C12N 15/85, 15/09
US CL :435/320.1, 172.3

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/320.1, 172.3

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, Biosis, World Patent Index.
Search terms: gene targeting, excision, integration

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Molecular and Cellular Biology, Vol. 11, No. 3, issued March 1991, Valancius et al., "Testing an 'in-out' targeting procedure for making subtle genomic modifications in mouse embryonic stem cells", pages 1402-1408. See entire article.	1, 4-7, 9
X – Y	Nature, Vol. 317, issued 19 September 1985, Smithies et al., "Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination", pages 230-234. See entire article.	1, 7 ——— 2,3,5,8-11,15-18
X – Y	Cell, Vol. 51, No. 6, issued 6 November 1987, Thomas et al., "Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells", pages 503-512. See entire article.	1,4,6,7,12, 13 ——— 2,3,5,9-11
X – Y	Nature, Vol. 345, issued 3 May 1990, DeChiara et al., "A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth fctor II gene disrupted by targeting", pages 78-80. See entire article.	12,13 ——— 14

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be part of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
16 August 1992

Date f mailing of the international search report
18AUG1992

Name and mailing address of the ISA/
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. NOT APPLICABLE

Authorized fficer
PHILIP W. CARTER
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*